

# Being Exceptional is Being Weak: Tonal Exceptions in San Miguel el Grande Mixtec

Eva Zimmermann  
*University of Leipzig*

## 1 Introduction

If a morpheme in a language employs a phonological alternation or shows some phonological behaviour that is unexpected given the phonological generalizations and rules established for the majority of morpheme alternations in this language, one can classify those morphemes as exceptions. One possible account for exceptionality in this sense is the assumption of morpheme-specific phonological constraints (e.g. Alderete, 2001; Pater, 2009; Finley, 2009) or phonological co-grammars (e.g. Orgun, 1996; Inkelas & Zoll, 2005, 2007): The morpheme behaves differently because it is subject to a different phonology. Or one could deny that this is still phonology and simply assume that the unexpected surface appearances or alternations of this exceptional morpheme are a consequence from different allomorphs listed in the lexicon (e.g. Paster, 2006, 2015; Bye, 2007). In contrast, I argue for a representational solution: That different morphemes behave differently with respect to the one phonology of the language is due to differences in their underlying representation. This is by no means a new claim and is in principle the same assumption defended in accounts based on, for example, underspecification (e.g. Archangeli, 1988; Inkelas, 1995; Krämer, 2000, 2001) or prosodic defectivity (e.g. Yearley, 1995; Stiebels & Wunderlich, 1999; Zimmermann, 2017c). The new perspective in this paper is that all phonological representations have a strength of activity or presence in an underlying representation. This is the proposal of Gradient Symbolic Representations made in Smolensky & Goldrick (2016) and Rosen (2016). In this paper, I strengthen the claim for Gradient Symbolic Representations in adding a case study of exceptional morphemes in the tonal phonology of San Miguel el Grande Mixtec. At the same time, I argue for a modification of the original Gradient Symbolic Representation account and for the concept of gradient activity in the output. This modification of the system predicts markedness-driven exceptions; a pattern that is borne out in the case study of San Miguel el Grande Mixtec.

This paper is structured as follows: In section 2, I present the data from San Miguel el Grande Mixtec. I will focus on two different instances of exceptional phonological behaviour of certain morphemes: Some morphemes are exceptional non-hosts for a tone if this creates a marked structure, and some morphemes show an unexpected alternation or allomorphy in surfacing with different tonal melodies in different phonological contexts. That the assumption of Gradient Symbolic Representations allows a unified account for these different exceptionality patterns is argued in section 3 where I first discuss the background assumptions on gradience (3.1) before I discuss the Harmonic Grammar account for the unexpected tone allomorphy (3.2) and the exceptional non-hosts (3.3). I conclude in section 4.

## 2 Data: Exceptions in the tonal phonology of MIG

San Miguel el Grande Mixtec (=MIG) is a variety of Mixtec (Otomanguean), spoken in Southern Mexico (Pike, 1944, 1948; Mak, 1950; McKendry, 2013). Most of the following data is taken from Mak (1950) and McKendry (2013). The language has the three level tones: H (=á), M (=a), and L (=à) and sequences of two tones are only possible on long vowels. This restriction follows straightforwardly under the assumption that the mora (=μ) is the tone-bearing unit (=TBU) in MIG and no true contour tones exist.

As nearly all Mixtec languages (Dürr, 1987; Hollenbach, 2003), MIG has ‘perturbing’ morphemes that

\* For valuable comments and discussion, I am grateful to the participants of the Leipzig Phonology Reading Group; especially to Daniel Gleim, Jochen Trommer, and Sören Worbs. I also want to thank the audiences at the mfm 2017 and AMP 2017. This research was supported by a DFG grant to the project ‘Featural Affixes: The Morphology of Phonological Features’ (TR 521/6–1).

change the tones of a following morpheme. In MIG, these tonal changes always involve the realization of an additional H on the initial TBU of a following morpheme<sup>1</sup> as can be seen in (1).<sup>2</sup> where perturbing morphemes are notated with a superscripted <sup>(H)</sup>. A comparison between perturbing morphemes and non-perturbing ones reveals that there is no systematic phonological or semantic property that sets them apart. Being a perturbing morpheme or not is hence a truly idiosyncratic property of morphemes.

(1) *MIG: Floating H's (Mak, 1950; McKendry, 2013)*

	Morpheme 1		Morpheme 2		Surface
a.	kəbð <sup>(H)</sup> ‘day’		biko ‘fiesta’		kəbð bíko M:83
b.	<sup>n</sup> deju <sup>(H)</sup> ‘food’		bàʔa ‘good’		<sup>n</sup> deju báʔa M:83
c.	k <sup>w</sup> aʔa <sup>(H)</sup> ‘many’		sùtʃí ‘children’		k <sup>w</sup> aʔà sútʃí M:83
d.	ʃíní <sup>(H)</sup> ‘head’		tʃíʔí ‘skunk’		ʃíní tʃíʔí McK:85
e.	nuù <sup>(H)</sup> ‘face’		nutʃí <sup>(H)</sup> ‘beans’		nuù núʃí McK:84
f.	βáá <sup>(H)</sup> EMPH		-tí <sup>(H)</sup> 3.ANIM		βáátí McK:92

Inside autosegmental phonology, a standard account for such a tone change is the assumption of underlyingly unassociated (or floating) tones in the representation of certain morphemes. Those floating tones then have to associate to a following morpheme and partially overwrite the underlying tone specification of this morpheme. However, several exceptional morphemes do not follow under such a simple autosegmental account (McKendry, 2013). One example are morphemes that are exceptional non-hosts for a floating H if the preceding morpheme sponsoring the H ends in an H-toned TBU. An example is the morpheme /-ðe/ 3.MHON as is shown in (2) (Pike, 1948:91). The examples in (2-a+b) show that this underlyingly M-toned morpheme becomes H-toned after a perturbing morpheme that ends in M or L – absolutely in line with what we saw in (1). Crucially, however, it remains M-toned if the preceding morpheme is perturbing and ends in an H-toned final TBU (2-c). A reasonable hypothesis is that this non non-realization of the floating H is due to a general OCP that bans two adjacent H's. However, as can be seen in the data in (2-d+e), floating H's are indeed associated on other non-exceptional morphemes to create such a marked HH-sequence. It is hence a truly idiosyncratic property of some morphemes that they are non-hosts for floating tones is this would create a marked OCP-violating structure.

(2) *Exceptional non-host for floating H (McKendry, 2013)*

	Morpheme 1		Morpheme 2		Surface
a.	nutʃí <sup>(H)</sup> ‘bean’		-ðe 3.MHON		nutʃiðé McK:92
b.	jee <sup>(H)</sup> ‘eat’		-ðe 3.MHON		jeeðé McK:104
c.	βáá <sup>(H)</sup> EMPH		-ðe 3.MHON		βááðe McK:92
d.	βáá <sup>(H)</sup> EMPH		-tí <sup>(H)</sup> 3.ANIM		βáátí McK:92
e.	ʃíní <sup>(H)</sup> ‘head’		tʃíʔí ‘skunk’		ʃíní tʃíʔí McK:85

Another example of exceptional tonal behaviour can be found for the two enclitics /jo/ 1.INCL and /ro/ 2 whose surface tones alternates between H, M, and L depending on the preceding morpheme. Examples for these morphemes with different tones are given in (3).

<sup>1</sup> This is a simplification. There are two exceptions to this that have received some theoretical discussion: bases with the tonal melody MH remain MH and ML bases of the shape (C)CVCV or CV<sub>1</sub>V<sub>2</sub> become MH. This deviation from realizing H on the initial TBU have been analysed as the result of a tonal metathesis rule (Goldsmith, 1990), special tone-status faithfulness (Tranel, 1995a,b), or tonal infixation (Zimmermann, 2016). The present discussion only focusses on the exceptional tonal behaviour of certain morphemes and is compatible with different ways to account for this non-local tone realization.

<sup>2</sup> Nasalization is a feature of morphemes in MIG (Marlett, 1992), notated as /CVCV/.

(3) *Surface realizations of /jo/ro/ (Pike, 1948; McKendry, 2013)*

Preceding morpheme			Surface		
a.	hinì	‘know’	hinìjò	L	McK:93
b.	sáʔa	‘make’	sáʔajó	H	McK:93
c.	kunu	‘run’	kunujó	H	P:90
d.	kee <sup>(H)</sup>	‘eat’	keero	M	P:91

A close look at the distribution of surface tones on these tones reveals that it is indeed completely phonologically predictable. The list in (4) summarizes all tones of these two morphemes together with their respective context of appearance. However, although the choice of tone is completely predictable from the tone of the preceding morpheme, there are no straightforward phonological principles or constraints deriving the tones from each other. In some contexts, the tone of /jo/ and /ro/ mirrors the tone of the preceding morpheme (C1, C3, C4, C6) but seems rather arbitrary in other contexts (C2, C5). Crucially, they also don’t seem to host a floating H after a perturbing morpheme (C3+C4)).

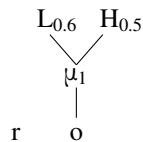
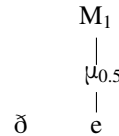
(4) *Tones of /jo/ro/: Phonological contexts (McKendry, 2013:93)*

following...	Surface
C1. L#	L
C2. H#	L
C3. L (H)#	L
C4. M (H)#	M
C5. M#	H
C6. H (H)#	H

**3 Analysis based on Gradient Symbolic Representations**

Both these pattern are unexpected given the general tonal phonology of the language: The tonal melodies on the morpheme /jo/ro/ seem to be an instance of phonologically conditioned but suppletive allomorphy (e.g. Paster, 2006) whereas the exceptional non-hosts are apparent exceptional non-undergoers for a regular phonological process. One strategy would hence be to assume that three different tonal allomorphs are stored for /jo/ro/ (Mascaró, 2007:e.g.) and that morpheme-specific phonological constraints or constructions are active for the exceptional non-hosts (e.g. Alderete, 2001; Pater, 2009; Finley, 2009). In contrast, I argue for a unified representational account where a simple difference in the representation of the /jo/ro/ and exceptional non-host morphemes predicts their exceptional behaviour from the regular phonology independently motivated for MIG. More concretely, I argue for the representations in (5) and (6): The morphemes /jo/ro/ are associated to two tones underlyingly that are both only weakly active and the exceptional non-hosts contain a  $\mu$  that is only weakly active, expressed by subscripted numerical activities smaller than 1.0.

The crucial theoretical ingredient in this account is hence the assumption of gradient activity that is introduced in more detail in the next section 3.1 before I turn to the concrete analysis of the /jo/ro/-allomorphy in section 3.2 and the exceptional non-hosts in section 3.3.

(5) *Representation: /jo/ro/*(6) *Representation: Exceptional non-host*

**3.1 Theoretical background: Gradient Symbolic Representations** The core assumption of Gradient Symbolic Representations as proposed in Smolensky & Goldrick (2016) is that phonological elements can have different degrees of presence in an underlying representation. Those activities are expressed as numerical activities where 1.0 is the default full activity. The underlying representation of a morpheme may,

however, also contain weakly active elements that are treated differently in the phonology and predict, for example, that these elements are not protected by faithfulness in the same way as fully active elements. This new dimension in which lexical representations might differ allows a unified explanation for various patterns of lexical exceptions in the phonology or allomorphy more general. Case studies assuming Gradient Symbolic Representations I am aware of cover things like the often-discussed liaison consonants in French (Smolensky & Goldrick, 2016), the semi-regularity of Japanese Rendaku (Rosen, 2016), allomorphy in Modern Hebrew (Faust & Smolensky, 2017), lexical accent in Lithuanian (Kushnir, 2017, 2018), lexical stress in Moses Columbian Salishan (Zimmermann, 2017a,b), or tone sandhi in Oku (Nformi & Worbs, 2017).

The specific formal implementation of the effects Gradient Symbolic Representations can have for the phonology are briefly illustrated with the toy example in (8). The grammatical computation is formally modelled in Harmonic Grammar where constraints are weighted, not ranked and the optimal candidate is the one with the best harmony score resulting from the sum of constraint violations times their respective weight (Legendre et al., 1990; Potts et al., 2010). The first crucial consequence from gradient activity in the input are gradient violations of faithfulness constraints. As can be seen in (8), deletion of the fully active obstruent  $t_1$  in candidate (8-b) results in a full violation of MAXS by -1.0, deletion of the partially active obstruent  $p_{0.5}$  in candidate (8-c), however, only violates it by -0.5.<sup>3</sup> DEPS is violated by -0.5 in (8-a+b) since the underlyingly weakly active  $p_{0.5}$  is neutralized to a fully active segment via insertion of activity.

- (7) a. MAXS: Assign violation X for any segmental activity X in the input that is not present in the output.  
 b. DEPS: Assign violation X for any segmental activity X present in the output but not in the input.
- (8) *Gradience and faithfulness violations*

$b_1 a_1 t_1 - p_{0.5}$	MAXS	DEPS	
	5	4	
Ⓐ a. batp		-0.5	-2
b. bap	-1	-0.5	-7
c. bat	-0.5		-2.5

In the original proposal in Smolensky & Goldrick (2016), all output elements have the full activity of 1.0: potentially weakly active input elements are hence neutralized to fully active ones in the output and different underlying activities only have consequences for evaluating faithfulness constraints. In contrast, it is now argued that elements may retain their (weak) activity in phonological output structure and that leveling to full activity for the phonetic interpretation happens at a later stage. The crucial consequence from this assumption is that the evaluation of markedness constraints is influenced by different activities as well: they are violated/satisfied to a weaker degree if they refer to weakly active elements. More concretely, if a markedness constraint \*A is violated by a structure A that has an activity smaller than 1, \*A is only violated by this number.<sup>4</sup> If, however, a constraint A! is only satisfied by a structure A that has an activity smaller than 1, this constraint is still violated by -1-(activity of the structure). Weakly active elements are hence not as bad a problem for constraints penalizing a certain structure and not as good a solution for constraints requiring a certain structure. This is illustrated with the extended toy example in (10). \*CC]<sub>σ</sub> penalizes consonant clusters and the fully active one in candidate (10-b) violates it more than the weakly active one in (10-a) whereas ONS! demands an onset consonant and the fully active onset consonant (10-b+c) perfectly satisfies it where as the weakly active one in (10-a) still violates it.

- (9) a. \*CC]<sub>σ</sub>: Assign violation X for a  $C_1 C_2$  in a syllable margin where X is the highest shared activity that both  $C_1$  and  $C_2$  have.  
 b. ONS!: Assign violation 1-X for every  $\sigma$  with an onset of activity X.

<sup>3</sup> Note that in contrast to this formalization here, MAX is a rewarding constraint in the original proposal in Smolensky & Goldrick (2016).

<sup>4</sup> In the following, the markedness constraints refer to the highest activity that is shared by all components of the marked structure. There are other possible formalizations of this but all of them predict the crucial asymmetry that weakly active elements are treated differently by markedness constraints.

(10) *Gradience and Markedness violations*

$b_{0.5}a_1t_1-p_{0.5}$	$*CC]_{\sigma}$	ONS!	DEP	MAX	
	4	4	2	2	
a. $b_{0.5}a_1t_1p_{0.5}$	-0.5	-0.5			-4
b. $b_1a_1t_1p_1$	-1		-1		-6
c. $b_1a_1t_1$			-0.5	-0.5	-2

**3.2 Allomorphy for /jo/ro/** The first part of the analysis is to show how the alternations of surfacing with either L, M, or H for the morphemes /jo/ro/ follows from a representation where these morphemes are assumed to be underlyingly associated to two weakly active tones:  $L_{0.6}$  and  $H_{0.5}$ . The first crucial observation is that realization of both these underlying tones results in a contour tone – but those are generally excluded in MIG. In the analysis, this follows from the (symbolic) high weight of 200 for  $*CONT$  (11-a). It is hence inherently impossible to realize the underlying tonal specification of these morphemes<sup>5</sup> as is shown in tableaux (12). The faithful candidate (12-a) is excluded by its violation of  $*CONT$ . The grammar thus has to choose between one of these two underlying tones. Given that the L has the slightly higher activity of 0.6 than the H that only has the activity of 0.5, realization of the former is preferred. This follows from the interaction of MAXT (11-b) preserving tonal input activity and SPEC (11-c) demanding that every TBU is associated to a tone. Non-realization of  $L_{0.6}$  in favor of realization of  $H_{0.5}$  (12-b) violates MAXT by -0.6 and also induces a -0.5 violation of SPEC, simply because that is the activity this tone lacks to a fully active tone  $T_1$  that would perfectly satisfy SPEC. Non-realization of  $H_{0.5}$  in favour of the more active tone  $L_{0.6}$  is better for both these constraints (12-c): MAXT is only violated by -0.5 and SPEC only by -0.4. This gradient difference in constraint violations makes the crucial difference for the harmony score and prefers optimal (12-c) over (12-b).

## (11)

- $*CONT$ : Assign violation X for every TBU associated to tones  $T_1$  and  $T_2$  where X is the highest activity that  $T_1$  and  $T_2$  share.
- MAXT: Assign violation X for any tonal activity X in the input that is not present in the output.
- SPEC: Assign violation 1-X for every TBU where X is the tonal activity associated to this TBU.

(12) *Preference for realizing  $L_{0.6}$* 

$H_1 + L_{0.6}H_{0.5}$ $\mu_1$	$*CONT$	SPEC	MAXT	
	200	70	10	
a. $H_1 L_{0.6} H_{0.5}$ $\mu_1$	-0.5			-100
b. $H_1 H_{0.5}$ $\mu_1$		-0.5	-0.6	-41
c. $H_1 L_{0.6}$ $\mu_1$		-0.4	-0.5	-33

This inherent preference for realization of  $L_{0.6}$  is overcome for bases ending in an M-toned  $\mu$  due to an additional markedness constraint (13) that penalizes every sequence of M followed by L. Realization of the in principle preferred L on a /jo/ro/-suffix as in (14-a) creates this marked sequence and hence induces an additional violation of  $*ML$ . Note that  $*ML$  is not fully violated in (13-a) but only by -0.6 since this is the activation of the L and hence the highest shared activity of the M and the L that form the marked sequence. This violation is still bad enough and realization of the less active  $H_{0.5}$  becomes optimal in such a context (14-b). The crucial weighting argument illustrated in (13) is hence that the weight of  $0.6x*ML$  is more than the weight of  $0.1xSPEC$  and  $0.1xMAXT$  taken together.

<sup>5</sup> This is in an interesting difference to the ‘blend structures’ in Smolensky & Goldrick (2016): For those, different gradiently active elements occupy one slot in the representation and their complementary distribution hence follows from the inherent nature of these special blend structures. In the present analysis, the complementary distribution of phonological elements competing for one position follows from standard phonological constraints independently motivated in the language.

- (13) \*ML: Assign violation X for every sequence of tone M followed by tone L where X is the highest activity that both share.
- (14) *Ban on ML makes  $L_{0.6}$  impossible*

$\begin{array}{c} M_1 \\   \\ \mu_1 \end{array} + \begin{array}{c} L_{0.6} \quad H_{0.5} \\ \diagdown \quad \diagup \\ \mu_1 \end{array}$		SPEC 70	*ML 15	MAXT 10	
a.	$\begin{array}{c} M_1 \quad L_{0.6} \\   \quad   \\ \mu_1 \quad \mu_1 \end{array}$	-0.4	-0.6	-0.5	-42
b.	$\begin{array}{c} M_1 \quad H_{0.5} \\   \quad   \\ \mu_1 \quad \mu_1 \end{array}$	-0.5		-0.6	-41

\*ML has not a particular high weight in MIG and expected sequences of ML in all contexts outside of /jo/ro/ surface faithfully because all other repair options to avoid ML are worse in these contexts.<sup>6</sup> Only if the grammar has to choose between the two tones of the /jo/ro/ allomorph, does this markedness preference emerge.

So far, we were only concerned with the choice between realizing one of the two tones to which /jo/ro/ are associated to underlyingly: Either  $L_{0.6}$  or  $H_{0.5}$ . However, both these strategies are not perfect to begin with since both still induce some violation of SPEC since the tones are not fully active. Since SPEC has a relatively high weight, providing the affixes with a fully active tone should always be better. At this point it is important to emphasize that DEPT is taken to have the high weight of 200 in MIG: This not only ensures that all weakly active tones remain weakly active in the output but of course also excludes insertion of a fully active epenthetic tone to avoid any SPEC-violations for /jo/ro/. As we will see below, most other strategies to provide /jo/ro/ with a fully active tone are excluded by independent constraints as well. However, we start the discussion with a tableau (16) that shows the one context where /jo/ro/ can indeed be associated to a fully active tone leaving their underlying weakly active tones unrealized. This context is a preceding morpheme ending in an M-toned or L-toned  $\mu$  followed by an additional unassociated H. Recall from (1) that the ‘perturbing’ morphemes in MIG are taken to be the ones that end in an unassociated or floating H (cf. 3.3 for more details on floating H-association). In such a context, spreading this underlying tone to /jo/ro/ (16-b) avoids any violation of SPEC and only induces a violation of \*LNGT penalizing tone spreading (15). Since even the partial violation of 0.4xSPEC has a higher weight than the violation of \*LNGT and 0.6xMAXT, this candidate has a better harmony score than the one realizing the underlying  $L_{0.6}$  tone (16-a).

- (15) \*LNGT: Assign violation X for every tone T that is associated to two TBU’s  $\tau_1$  and  $\tau_2$  where X is the highest activity T,  $\tau_1$ , and  $\tau_2$  share.
- (16) *Spreading of a stem-tone*

$\begin{array}{c} M_1 \quad H_1 + L_{0.6} \quad H_{0.5} \\   \quad \quad \diagdown \quad \diagup \\ \mu_1 \quad \quad \mu_1 \end{array}$		SPEC 70	*LNGT 15	MAXT 10	
a.	$\begin{array}{c} M_1 \quad H_1 \quad L_{0.6} \\   \quad \quad   \\ \mu_1 \quad \quad \mu_1 \end{array}$	-0.4		-0.5	-33
b.	$\begin{array}{c} M_1 \quad H_1 \\   \quad \quad \diagdown \\ \mu_1 \quad \quad \mu_1 \end{array}$		-1	-1.1	-26

Another obvious candidate in (16) is one that simply associates the floating tone of the preceding morpheme. Association of such a floating tone to /jo/ro/ solves two problems: It ensures that the floating tone has a host

<sup>6</sup> Two obvious strategies are simply deleting one of the tones and leaving the  $\mu$  unspecified or deleting one of the tones and providing a new one via spreading. The former option is excluded since the combined weight of MAXT + SPEC is (way) higher than the weight of \*ML and the latter since the combined weight of MAXT + \*LNGT (cf. (15)) is higher than the weight of \*ML.

and that /jo/ro/ are associated to a fully active tone. It is assumed that the constraint (17) excludes such a strategy in general: Floating H's can never associate to a  $\mu$  that was already underlyingly associated to a H.<sup>7</sup> Since the effect of perturbing morphemes can only be detected for  $\mu$ 's underlyingly specified for M or L, this constraint is surface-true and taken to have a very high weight of 200 in MIG. Its effect is shown in tableau (18) that simply adds the relevant candidate (18-b) to tableau (16).

- (17) DEP-H: Assign a violation 1 for every new association between a H that is unassociated in the input and a TBU if the TBU was underlyingly associated to a H.

- (18) *No association of a floating H*

	$M_1 H_1 + L_{0.6} H_{0.5}$ $\mu_1$	DEP-H 200	SPEC 70	*LNGT 15	MAXT 10	
a.	$M_1 H_1$ $\mu_1$			-1	-1.1	-26
b.	$M_1 H_1$ $\mu_1$	-1			-1.1	-211

A second important constraint penalizing an alternative strategy to provide /jo/ro/ with a fully active tone is DEP-F (19). It is a positional faithfulness constraint penalizing new associations of morpheme-final tones.<sup>8</sup> Its effect is simple: it bans spreading of any tone that is the rightmost in the tonal melody of a morpheme and hence of any tone that is not followed by a floating H. Crucially, the optimal candidate in (16) does not violate DEP-F since the spreading tone is not the final tone of this morpheme: A floating tone hence shields the tones preceding it inside the same morpheme from DEP-F violations. That DEP-F blocks spreading of a preceding tone if this preceding morpheme does not end in a floating H is shown in tableau (20).<sup>9</sup>

- (19) DEP-F: Assign a violation 1 for every epenthetic association between a TBU and a tone that is morpheme-final.

- (20) *No spreading for stem-final tones*

	$M_1 + L_{0.6} H_{0.5}$ $\mu_1$	SPEC 70	DEP-F 19	*ML 15	*LNGT 15	MAXT 10	
a.	$M_1 L_{0.6}$ $\mu_1$	-0.4		-0.6		-0.5	-42
b.	$M_1 H_{0.5}$ $\mu_1$	-0.5				-0.6	-41
c.	$M_1$ $\mu_1$		-1		-1	-1.1	-45

The analysis of the tone allomorphy for /jo/ro/ can hence be summarized as in (21), referring back to the six contexts given in (4). The two morphemes are taken to be underlyingly associated to two weakly active tones. Realization of these is sub-optimal from the perspective of SPEC and spreading of a preceding tone is in principle preferred. This, however is only possible if this tone is not the morpheme-final one: in essence,

<sup>7</sup> This is interpreted as a context-sensitive version of a DEP-ASSOCIATIONLINE constraint and as such not scaled for activation (cf. Smolensky & Goldrick (2016) for the same assumption that all faithfulness constraints but MAX and DEP are not gradient (p.17)).

<sup>8</sup> Cf. Krämer (2003); Barnes (2008) on final syllable prominence effects.

<sup>9</sup> DEP-F is of course also violated if a floating H associates to /jo/ro/ as in candidate b. in (18) since floating H's in MIG are always morpheme-final. Since association of the floating H gets rid of a H→ $\mu$ -violation (cf. 3.3), only the combined weight of DEP-F and DEP-H excludes such a strategy.

this means that only tones that are followed by a floating H can ever spread to /jo/ro/. This predicts the fact that /jo/ro/ mirrors the tone of the final  $\mu$  of a preceding perturbing morpheme (21-A). Note that DEP<sub>-H</sub> is only formulated for underlyingly floating tones: the spreading in (21-C6) is hence not excluded. If such a spreading of a preceding tone is impossible (21-B), /jo/ro/ have to content themselves with their weakly activated underlying tones. The choice between those is in principle made in favor of more active L<sub>0.6</sub> (21-C1+C2) that is only blocked if the preceding morpheme ends in an M-toned  $\mu$  (21-C5). The complex allomorphy on /jo/ro/ hence indeed falls out from a single underlying representation under the assumption of gradient activity for phonological elements.

(21) *Summary of the analysis*

A. Spreading of non-final stem tone possible				B. Spreading of stem-final tone impossible			
C3.	L (H)#	L	$\begin{array}{c} L_1 \ H_1 \\   \quad \diagdown \\ \mu_1 \quad \mu_1 \end{array}$	C1.	L#	L	$\begin{array}{c} L_1 \quad L_{0.6} \\   \quad \quad   \\ \mu_1 \quad \mu_1 \end{array}$
C4.	M (H)#	M	$\begin{array}{c} M_1 \ H_1 \\   \quad \diagdown \\ \mu_1 \quad \mu_1 \end{array}$	C2.	H#	L	$\begin{array}{c} H_1 \quad L_{0.6} \\   \quad \quad   \\ \mu_1 \quad \mu_1 \end{array}$
C6.	H (H)#	H	$\begin{array}{c} H_1 \ H_1 \\   \quad \diagdown \\ \mu_1 \quad \mu_1 \end{array}$	C5.	M#	H	$\begin{array}{c} M_1 \quad H_{0.5} \\   \quad \quad   \\ \mu_1 \quad \mu_1 \end{array}$

**3.3 Exceptional non-hosts** We already saw some floating tones in the analysis of /jo/ro/ in the last subsection: Those are simply taken to be unassociated H's at the end of the tonal melody of certain morphemes in MIG. In this subsection we now turn to the concrete analysis of these floating tones in order to explain the behaviour of the exceptional non-hosts. Recall that the generalization about the perturbing morphemes was that their tone was (nearly) always realized on the initial  $\mu$  of a following morpheme except for an idiosyncratic class of morphemes that kept their underlying tonal specification if a new sequence of two adjacent H-toned TBU's would be created otherwise.

A formal account for these exceptional non-hosts has to begin with an account of why floating H's are realized in the first place. The OT-implementation for such a standard instance of tonal overwriting is based on the constraints in (22) and follows the account for floating feature association in Wolf (2007). The two crucial constraints driving overwriting are (22-a) and (22-b) demanding that all elements in the output must be associated and that an underlyingly floating element may not be deleted. In addition, \*CONT is again crucial that excludes contour tones in MIG with its high weight of 200 (cf. (11)).

- (22) a. MAXFL: Assign violation X for any activity X of a floating tone in the input that is not present in the output.  
 b. H→ $\mu$ : Assign 1-X violations for every H-tone where X is the activity of TBU's this H is associated to.

The weightings of these constraints predicting H-overwriting are illustrated in tableau (23) where a morpheme ending in an L-toned  $\mu$  also contains a floating H and precedes a morpheme that begins with an M-toned  $\mu$ . The faithful candidate (23-a) fatally violates H→ $\mu$  since the H is not associated to any  $\mu$ . Deletion of this H avoids the violation of H→ $\mu$  but induces a violation of MAXFL (23-b). Associating the H to the following morpheme as in (23-c) also avoids the H→ $\mu$  violation but creates a contour tone that is generally illicit in MIG and penalized by \*CONT. Both MAXFL and \*CONT are never violated in MIG; mirrored in their very high weight of 200. The contour tone in (23-c) is avoided in the 'overwriting' candidate (23-d) that associates the H to a following  $\mu$  and deletes the tone this  $\mu$  was underlyingly associated to. It violates the general MAXT constraint that has a lower weight than H→ $\mu$ . This candidate hence emerges as the winner. A final relevant candidate not shown in (23) is one that overwrites the final tone of the first morpheme and not the initial one of the following. This candidate is excluded by ALTERNATION (van Oostendorp, 2007:39) that penalizes new associations within one morpheme and has a high weight of 200. A floating tone hence has to associate to another morpheme, it can never associate to TBU's of its own morpheme.

Note that this tableau also includes DEP<sub>-F</sub> whose weight was determined in the preceding analysis of the /jo/ro/-allomorphy. The constraint in fact makes overwriting worse since a floating tone is always the final one of a morpheme and its realization hence always implies a DEP<sub>-F</sub> violation. Since the combined weight



of DEP<sub>-F</sub> + MAXT is still lower than the weight of H→μ, overwriting is nevertheless predicted.

(23)

$\begin{array}{c} L_1 \quad H_1 \\   \quad   \\ \mu_1 \end{array} + \begin{array}{c} M_1 \\   \\ \mu_1 \end{array}$	MAXFL	*CONT	H→μ	DEP <sub>-F</sub>	MAXT	
	200	200	60	19	10	
a. $\begin{array}{c} L_1 \quad H_1 \quad M_1 \\   \quad \quad   \\ \mu_1 \end{array}$			-1			-60
b. $\begin{array}{c} L_1 \quad \quad M_1 \\   \quad \quad   \\ \mu_1 \end{array}$	-1				-1	-210
c. $\begin{array}{c} L_1 \quad H_1 \quad M_1 \\   \quad \quad   \\ \mu_1 \end{array}$		-1		-1		-219
d. $\begin{array}{c} L_1 \quad \quad H_1 \\   \quad \quad   \\ \mu_1 \end{array}$				-1	-1	-29

Turning now to the exceptional non-hosts for floating tones, a crucial markedness constraint is the OCP penalizing H's on adjacent TBU's and hence μ's in MIG (24).

(24) OCP: Assign X violations for every pair of adjacent TBU's τ<sub>1</sub> and τ<sub>2</sub> associated to different H's if X is the highest activity that τ<sub>1</sub> and τ<sub>2</sub> share.

The weight of the OCP in MIG is taken to be relatively low and crucially so low that even the combined weight of the OCP, DEP<sub>-F</sub>, and MAXT is lower than the weight of H→μ. This predicts that a floating H overwrites and is hence realized even if it follows a morpheme ending in a H-toned μ. Realization of a floating H in such a context (25-b) hence creates an additional violation of the OCP (in addition to the violations of DEP<sub>-F</sub> and MAXT that every tone overwriting induces) but the resulting harmony score is still better than the one resulting from leaving the tone floating (25-a). Violations of the OCP are hence in principle tolerated since realization of a floating tone is so important.

(25) *Floating H associates to a strong host: OCP irrelevant*

$\begin{array}{c} H_1 \quad H_1 \\   \quad   \\ \mu_1 \end{array} + \begin{array}{c} L_1 \\   \\ \mu_1 \end{array}$	MAXFL	H→μ	DEP <sub>-F</sub>	OCP	MAXT	
	200	60	19	10	10	
a. $\begin{array}{c} H_1 \quad H_1 \quad L_1 \\   \quad \quad   \\ \mu_1 \end{array}$		-1				-60
b. $\begin{array}{c} H_1 \quad \quad H_1 \\   \quad \quad   \\ \mu_1 \end{array}$			-1	-1	-1	-39

With this background in place, we can now turn to the morphemes that are exceptional non-hosts for floating H's. The crucial assumption is that those end in a μ that is only partially active and has a gradient activity of 0.5. Since the μ is taken to be the TBU in MIG, those morphemes hence contain a somewhat defective TBU and any association to this μ is not a perfect strategy for any tone. More concretely, any H that is associated to this weakly active μ violates H→μ by -0.5. This follows because H→μ is a positive constraint demanding that every tone must be associated to a μ – it is hence only satisfied if every tone is associated to a fully active μ<sub>1</sub> and any activity that is lacking to this optimal structure is a violation. That this partial solution to the problem of floating H's is in principle sufficient is shown in tableau (26) where a base ending in an L-toned μ contains an additional floating H and precedes a morpheme that starts with a weakly active μ<sub>0.5</sub>. Leaving this tone floating violates H→μ (26-a) whereas overwriting as in (26-b) induces not only a MAXT

and a DEP<sub>-F</sub> violation (=given that the floating H is the final tone of a morpheme), but still violates H→μ by -0.5. However, the weights of these violations are not yet enough to overcome the one violation of H→μ and (26-b) is hence optimal. Even solving the H→μ-problem only partially is hence still better than not solving it at all. Crucially, the OCP is irrelevant in this case since no underlying H precedes the floating H.

(26) *Floating H associates to a weak host: No OCP-violation*

$\begin{array}{c} M_1 \quad H_1 \\   \\ \mu_1 \end{array} + \begin{array}{c} M_1 \\   \\ \mu_{0.5} \end{array}$		MAXFL	H→μ	DEP <sub>-F</sub>	OCP	MAXT	
		200	60	19	10	10	
a.	$\begin{array}{c} M_1 \quad H_1 \quad M_1 \\   \quad \quad   \\ \mu_1 \quad \quad \mu_{0.5} \end{array}$		-1				-60
b.	$\begin{array}{c} M_1 \quad \quad H_1 \\   \quad \quad   \\ \mu_1 \quad \quad \mu_{0.5} \end{array}$		-0.5	-1		-1	-59

This picture shifts dramatically, however, if the morpheme sponsoring the floating H ends in H and realization of the floating tone hence induces an additional OCP violation. This can be seen in tableau (27): non-realization of the floating H (27-a) induces a full violation of H→μ whereas association to the following μ still violates H→μ by -0.5, has a MAXT, a DEP<sub>-F</sub>, and crucially also an OCP violation. The latter is now what crosses the threshold in favour of the floating candidate (27-a) that is optimal in this context.

(27) *Floating H association blocked for weak hosts: Fatal OCP-violation*

$\begin{array}{c} H_1 \quad H_1 \\   \\ \mu_1 \end{array} + \begin{array}{c} M_1 \\   \\ \mu_{0.5} \end{array}$		MAXFL	H→μ	DEP <sub>-F</sub>	OCP	MAXT	
		200	60	19	10	10	
a.	$\begin{array}{c} H_1 \quad H_1 \quad M_1 \\   \quad \quad   \\ \mu_1 \quad \quad \mu_{0.5} \end{array}$		-1				-60
b.	$\begin{array}{c} H_1 \quad \quad H_1 \\   \quad \quad   \\ \mu_1 \quad \quad \mu_{0.5} \end{array}$		-0.5	-1	-1	-1	-69

This crucial weighting argument for the exceptional non-hosts is summarized in (28). They show a clear threshold effect only possible under gradient representations and hence gradient constraint violations: Whereas H→μ is more important than DEP<sub>-F</sub>, OCP, and MAXT (28-a), the weighting relation is reversed for half a H→μ violation (28-c).

(28) *Weak Hosts: Threshold effects*

	Weight of	is greater than	the weight of	Cf.
a.	H→μ	≫	DEP <sub>-F</sub> + OCP + MAXT	(25)
b.	0.5xH→μ	≫	DEP <sub>-F</sub> + MAXT	(26)
c.	DEP <sub>-F</sub> + OCP + MAXT	≫	0.5xH→μ	(27)

For this account, output structures must crucially be gradient as well since the triggering constraints for the asymmetric behaviour of certain morphemes are markedness constraints; most crucially H→μ. Under the original proposal in Smolensky & Goldrick (2016), such gradient violations of markedness constraints are impossible and exceptionality patterns as in MIG cannot be predicted.

## 4 Conclusion

In this paper, an analysis of two seemingly unrelated exceptionality patterns in the tonal phonology of San Miguel el Grande Mixtec was proposed. The exceptional non-hosts simply looked like an instance of an exceptional non-undergoer for an otherwise regular process in a language and the tonal alternations on /jo/ro/ like a classical case of phonologically conditioned suppletive allomorphy. The main claim in this paper is that exceptionality patterns like that are not unrelated at all since a unified account for them is possible once Gradient Symbolic Representations are taken into account. Underlying phonological elements can be inherently more marked (=the weak tones on the /jo/ro/) than other seemingly identical phonological elements or can be not as bad a problem or not as good a solution (=the weak non-hosts for floating tones) as other seemingly identical phonological elements. Crucially, the analyses relied on gradience in the output and hence a representational strength difference that has consequences for the evaluation of markedness constraints. This sets the account apart from the original proposal of Gradient Symbolic Representations in Smolensky & Goldrick (2016); Rosen (2016). Apart from this new representational assumptions of gradient activity, the analysis is based primarily on standard constraints on tone structures (cf., for example, Yip, 2002) – the assumption of Gradient Symbolic Representations hence allows to predict a wide variety of apparent exceptionality patterns from a single and rather simple phonology.

## References

- Alderete, John (2001). *Morphologically governed accent in Optimality Theory*. Routledge, New York.
- Archangeli, Diana (1988). Aspects of underspecification theory. *Phonology* 5, 183–207.
- Barnes, Jonathan (2008). *Strength and Weakness at the Interface : Positional Neutralization in Phonetics and Phonology*. de Gruyter Mouton, Berlin, Boston.
- Bye, Patrick (2007). Allomorphy selection, not optimization. Blaho, Sylvia, Patrick Bye & Martin Krämer (eds.), *Freedom of Analysis?*, Mouton, 63–91.
- Dürr, Michael (1987). A preliminary reconstruction of the Proto-Mixtec tonal system. *Indiana* 11, 19–61.
- Faust, Noam & Paul Smolensky (2017). Activity as an alternative to autosegmental association. talk given at mfm 25, 27th May, 2017.
- Finley, Sara (2009). Morphemic harmony as featural correspondence. *Lingua* 119, 478–501.
- Goldsmith, John (ed.) (1990). *Autosegmental and Metrical Phonology*. Oxford: Blackwell.
- Hollenbach, Barbara (2003). The historical source of an irregular Mixtec tone-sandhi pattern. Wise, Mary Ruth, Thomas Headland & Ruth Brend (eds.), *Language and life: essays in memory of Kenneth L. Pike*, SIL international, Dallas, 535–552.
- Inkelas, Sharon (1995). The consequences of optimization for underspecification. *NELS* 25, 287–302.
- Inkelas, Sharon & Cheryl Zoll (2005). *Reduplication: Doubling in Morphology*. Cambridge University Press, Cambridge.
- Inkelas, Sharon & Cheryl Zoll (2007). Is grammar dependence real? A comparison between cophological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45:1, 133–171.
- Krämer, Martin (2000). Voicing alternations and underlying representations: the case of Breton. *Lingua* 9, 639–663.
- Krämer, Martin (2001). Yucatec Maya vowel alternations - harmony as syntagmatic identity. *Zeitschrift für Sprachwissenschaft* 20, 175–217.
- Krämer, Martin (2003). What is wrong with the right side? edge (a)symmetries in phonology and morphology. Unpublished paper. University of Ulster. Available on Rutgers Optimality Archive, ROA-576.
- Kushnir, Yuriy (2017). Accent strength in Lithuanian. talk, given at the workshop on Strength in Grammar, Leipzig, November 12, 2017.
- Kushnir, Yuriy (2018). Accent strength in Lithuanian prosody. talk, presented at OCP 15, January 14, 2018.
- Legendre, Geraldine, Yoshiro Miyata & Paul Smolensky (1990). Harmonic grammar – a formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. *Proceedings of the 12th annual conference of the cognitive science society* 388–395.
- Mak, Cornelia (1950). A unique tone perturbation in Mixteco. *International Journal of American Linguistics* 16, 82–86.
- Marlett, Stephen A. (1992). Nasalization in Mixtec languages. *International Journal of American Linguistics* 58, 425–435.
- Mascaró, Joan (2007). External allomorphy and lexical representation. *Linguistic Inquiry* 38, 715–735.
- McKendry, Inga (2013). *Tonal Association, Prominence and Prosodic Structure in South-Eastern Nochixtlán Mixtec*. Ph.D. thesis, University of Edinburgh.
- Nformi, Jude & Sören Worbs (2017). Gradient tones obviate floating features in Oku tone sandhi. talk, given at the workshop on Strength in Grammar, Leipzig, November 10, 2017.
- van Oostendorp, Marc (2007). Derived environment effects and consistency of exponence. Blaho, Sylvia, Patrik Bye & Martin Krämer (eds.), *Freedom of Analysis?*, Mouton, Berlin, 123–148.

- Orgun, Cemil Orhan (1996). *Sign-based Morphology and Phonology with Special Attention to Optimality Theory*. Ph.D. thesis, UC Berkeley.
- Paster, Mary (2006). *Phonological Conditions on Affixation*. Ph.D. thesis, University Of California, Berkeley.
- Paster, Mary (2015). Phonologically conditioned suppletive allomorphy: cross-linguistic results and theoretical consequences. Bonet, Eulàlia, Maria-Rosa Lloret & Joan Mascaró (eds.), *Understanding Allomorphy: Perspectives from Optimality Theory*, Equinox, 218–253.
- Pater, Joe (2009). Morpheme-specific phonology: Constraint indexation and inconsistency resolution. Parker, Steve (ed.), *Phonological Argumentation: Essays on Evidence and Motivation*, Equinox, London, 123–154.
- Pike, Kenneth L. (1944). Analysis of a Mixteco text. *International Journal of American Linguistics* 10, 113–138.
- Pike, Kenneth L. (1948). *Tone languages*. University of Michigan Press, Ann Arbor, MI.
- Potts, Christopher, Joe Pater, Karen Jesney, Rajesh Bhatt & Michael Becker (2010). Harmonic grammar with linear programming: From linear systems to linguistic typology. *Phonology* 77–117.
- Rosen, Eric (2016). Predicting the unpredictable: Capturing the apparent semi-regularity of rendaku voicing in Japanese through harmonic grammar. Clem, Emily, Virginia Dawson, Alice Shen, Amalia Horan Skilton, Geoff Bacon, Andrew Cheng & Erik Hans Maier (eds.), *Proceedings of BLS 42*, Berkeley Linguistic Society, 235–249.
- Smolensky, Paul & Matthew Goldrick (2016). Gradient symbolic representations in grammar: The case of French Liaison. *ROA* 1286.
- Stiebels, Barbara & Dieter Wunderlich (1999). Second stems in Hungarian nouns. *The Linguistic Review* 16, 253–294.
- Tranel, Bernard (1995a). On the status of universal association conventions: Evidence from Mixteco. Ahlers, J., L. Bilmes, J. Guenter, B. Kaisse & J. Namkung (eds.), *Proceedings of BLS 21*, 299–312.
- Tranel, Bernard (1995b). Rules vs. constraints: a case study. *ROA*-72.
- Wolf, Matthew (2007). For an autosegmental theory of mutation. Bateman, Leah, Michael O’Keefe, Ehren Reilly, & Adam Werle (eds.), *UMOP 32: Papers in Optimality Theory III*, GLSA, Amherst, MA, 315–404.
- Yearley, Jennifer (1995). Jer vowels in Russian. Beckman, Jill, Laura Walsh Dickey & Suzanne Urbanczyk (eds.), *Papers in Optimality Theory*, GLSA Publications, Amherst, Mass., 533–571.
- Yip, Moira (2002). *Tone*. Cambridge University Press.
- Zimmermann, Eva (2016). The power of a single representation: Morphological tone and allomorphy. *Morphology, special issue on allomorphy* 26, 269–294.
- Zimmermann, Eva (2017a). Being (slightly) stronger: Lexical stress in Moses Columbian Salish . talk, given at the workshop on Strength in Grammar, Leipzig, November 11, 2017.
- Zimmermann, Eva (2017b). Gradient symbolic representations in the output: A typology of lexical exceptions. talk, given at NELS 48, Reykjavik, October 29, 2017.
- Zimmermann, Eva (2017c). *Morphological Length and Prosodically Defective Morphemes*. Oxford University Press, Oxford.